Executive Summary

I. Introduction

Historical Background

The acute oral toxicity test in rodents is a critical step in defining the toxicity of a test material for the purpose of hazard classification and labeling. The acute oral toxicity test is designed to determine adverse effects and to estimate the dose level that is expected to kill 50% of the test population (i.e., the LD50).

A procedure for calculating the oral LD50 was first described by Trevan in 1927. This procedure has been used as a benchmark for comparing the toxicity of chemicals. The original method often used 50 animals or more. In 1981, the Organisation for Economic Co-operation and Development (OECD) adopted a test guideline (TG 401) for acute oral toxicity that estimated the LD50, and in some cases, the slope and confidence interval (CI). OECD TG 401 has become the traditional acute oral toxicity test. The test guideline was revised in 1987 to utilize three dose groups of five rats of one sex, with confirmation in the other sex using one group of five rats. In the absence of a range-finding study, this revision reduced the minimum number of animals used in the traditional acute oral toxicity test from 30 to 20.

In a continuing attempt to improve the estimate of acute toxicity while reducing the number of animals used per test, three alternative test methods were developed and implemented as additional OECD Guidelines for acute toxicity. These three tests are the Fixed Dose Procedure (FDP, TG 420), the Acute Toxic Class Method (ATCM, TG 423), and the Up-and-Down Procedure (UDP, TG 425).

U.S. EPA Request for Review of a Revised UDP

The U.S. Environmental Protection Agency (EPA) requested the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) organize an independent scientific peer review evaluation of the validation status of a revised Up-and-Down Procedure (UDP). The U.S. EPA forwarded the proposed "Acute Oral Toxicity: Modified Up-and-Down Procedure (Revised UDP)" to ICCVAM in April 2000. An independent peer review panel (Panel) was convened on July 25, 2000 to evaluate the proposed tests based on ICCVAM validation and regulatory acceptance criteria (NIEHS, 1997). An earlier version of the UDP test method had been adopted by the OECD TG Program in 1998 (TG 425). The revised UDP was proposed as an alternative to the existing conventional LD50 test (OECD TG 401, 1987; U.S. EPA 870.1100, 1998) used to assess the acute oral toxicity of chemicals. The U.S. EPA subsequently determined it was necessary to revise the UDP. The revisions were needed to 1) conform to a newly harmonised global hazard classification scheme for acute toxicity (OECD, 1998b; updated OECD, 2001); and 2) to incorporate changes to ensure the regulatory and testing needs would be met using the revised UDP prior to the OECD's proposed deletion of the TG 401 (OECD, 1987).

Components of the Revised UDP Test Method

The revised UDP test method submitted to ICCVAM in April 2000 included three components:

a) Primary Test, which provided an improved estimate of acute oral toxicity with a reduction in the number of animals used when compared to TG 401 and the existing TG 425;
b) Limit Test for substances anticipated to have minimal toxicity; and
c) Supplemental Test to determine the slope and confidence interval (CI) for the dose-response curve.

The Panel congratulates the agencies of the United States and the OECD for moving forward with the sequential testing of animals, as was achieved with the adoption of OECD TG 425 and in the proposed revision. Also, the development

team for the revised UDP demonstrated a comprehensive understanding of the statistical issues involved and is to be commended for the effort that went into revising the UDP Guideline.

In the revised UDP Primary Test, one animal is orally administered an appropriate dose (with 175 mg/kg as the default starting dose) and observed for up to 14 days. If the animal is alive at 48 hours after treatment, a second animal is orally administered a preset higher dose (0.5 log spacing by default). If the first animal dies, then the second animal is dosed at a preset lower dose (0.5 log spacing by default). Dosing stops when one of three stopping criteria is satisfied, with as few as six, but not more than 15 animals used per test.

In the revised UDP Limit Test, one animal is dosed at the limit dose (2000 or 5000 mg/kg). If the animal dies, the UDP Primary Test is conducted. If the animal lives, two more animals are dosed concurrently at the limit dose. If both of these animals live (i.e., three animals have survived), the UDP Limit Test is stopped. If one or both of the two animals die, additional animals are dosed sequentially at the limit dose until either three animals have survived or three animals have died (i.e., the maximum number of animals tested is five). If three animals survive, the LD50 is above the limit dose. Conversely, if three animals die, the LD50 is below the limit dose level.

In the UDP Supplemental Test for determining the slope and CI, three treatment schedules at increasing dose levels are initiated, each at a dose level that is a factor of 10- to 30-fold below the estimated LD50 obtained in the UDP Primary Test. Dosing continues in each sequence until an animal dies. All data, including data obtained in the UDP Primary Test, are then considered in a statistical model that estimates the slope and CI.

II. ICCVAM Independent Scientific Peer Review, July 25, 2000 Peer Review Meeting

In a public session on July 25, 2000, an international independent scientific peer review panel (Panel) met to evaluate the validation status of the revised UDP (*Federal Register*, NIEHS, 2000a, 2000b). The Panel was charged with evaluating the extent to which established ICCVAM validation and acceptance criteria had been addressed, and subsequently developing conclusions regarding the usefulness and limitations of the UDP. Evaluation of the Revised UDP was divided into four sections:

1. General Considerations for the Revised UDP Protocol;
2. Revised UDP Primary Test;
3. Revised UDP Limit Test; and
4. UDP Supplemental Test.

The Panel was also asked to respond to the following questions for each of the three tests:

- Has the revised UDP been evaluated sufficiently, and is its performance satisfactory to support its adoption as a substitute for the currently accepted UDP (OECD TG 425), and as a substitute for the conventional LD50 test for acute oral toxicity (U.S. EPA OPPTS 870.1100; OECD TG 401)?

- With respect to animal welfare, does the revised UDP adequately consider and incorporate where scientifically feasible, procedures to refine, reduce, and/or replace animal use?

In response to these questions, the Panel concluded the following:

1. The performance of the revised UDP Primary Test is satisfactory and exceeds the performance of OECD TG 401 in providing, with fewer animals, both an improved estimate of the LD50 for the purpose of hazard classification and more accurate information on acute toxicity. In particular, the use of 0.5 log units for dose spacing is reasonable and appropriate based on experience and the results of computer simulations. Three disadvantages of the revised UDP Primary Test recognized by the Panel were: a) the increased length of time needed to conduct a study; b) the increased costs per test material evaluated; and c) the increased complexity of the protocol.
2. The revised UDP Limit Test at 2000 or 5000 mg/kg is expected to perform as well as or

better than the Limit Test in OECD TG 401, with a reduction in the number of animals needed to conduct a test.

3. The UDP Supplemental Test for slope and CI was not recommended for adoption. The Panel was unable to evaluate the utility of the test because sufficient information regarding the use of the resulting data was not provided. As a consequence, any impact on animal use was not assessed.
4. The revised UDP Primary Test and the revised UDP Limit Test will reduce the number of animals used, but will not replace the use of animals. The Panel could not reach a consensus on the overall issue of refinement. However, the OECD Guidance Document on the Recognition, Assessment, and Use of Clinical Signs as Humane Endpoints for Experimental Animals used in Safety Evaluation (OECD, 2000a), referenced in the revised UDP Guideline, provides an element of refinement.

The recommendations of the Panel for the revised UDP follow. Additional information can be found in the appropriate sections of this report.

General Considerations

With regard to general protocol and UDP Guideline-related issues, the Panel recommended the following:

- The use of either sex (all males or all females) should be permitted unless information is available suggesting that one sex is more sensitive;
- The use of constant volume or constant concentration of the test material during administration should be allowed;
- All reference to littermates should be excluded from the U.S. EPA Revised UDP Guideline;
- Animals of 8 to 12 weeks of age should be used;
- Individual animal body weights on the day of dosing should be within 20% of the mean body weight for all animals dosed throughout the study;
- Additional guidance detailing how to use all pre-start data (e.g., *in vitro* test results, physical and chemical properties) should be provided in the Guideline;
- The overall usefulness of information (e.g., clinical signs, time course of effects, target organs, pathology, etc.) gained beyond the LD50 should be emphasized in the UDP Guideline; and
- The Guideline should be reorganized to improve clarity.

UDP Primary Test

With regard to the revised UDP Primary Test, the Panel recommended the following:

- The scientific basis should be presented in the Revised UDP Guideline;
- Guidance for when to use the UDP Primary Test should be included in the Guideline;
- Additional guidance on the starting rule and a justification of the default starting dose of 175 mg/kg should be discussed in the Guideline;
- An improved description of stopping rule #3 should be included in the Guideline;
- User-friendly, validated software for test use or access to such software should be provided; and
- A practicability evaluation should be conducted (an appropriate working group should consider the design of this evaluation).

UDP Limit Test

With regard to the revised UDP Limit Test, the Panel recommended:

- The scientific basis and rationale should be added to the Revised UDP Guideline; and
- Additional discussion of how and where the revised UDP Limit Test is integrated into the strategy of hazard or safety assessment should be included in the Guideline (a flow chart with decision criteria covering the complete testing scheme might be an efficient way to attain this goal).

UDP Supplemental Test

With regard to the UDP Supplemental Test, the Panel recommended:

- a more clearly defined purpose of how the slope and CI are used for human and environmental risk assessment should be included in the Revised UDP Guideline; and
- Consideration should be given as to whether the slope and CI are the most appropriate parameters for risk assessment or whether risk assessment needs can be addressed more

directly. For example, if estimates of points on the dose-response curve well below the median lethal dose are needed in environmental risk assessment, more efficient methods should be considered.

Revisions to the UDP in response to the July 25, 2000 Panel Report

Based on the Panel's conclusions and recommendations from July 25, 2000, the UDP Technical Task Force revised the UDP test method guideline as follows:

- Revisions recommended by the Panel were incorporated into the proposed UDP Primary and Limit Tests;
- The UDP Supplemental Test to determine the slope of the dose-response curve was deleted;
- A procedure was added (for use with the Primary Test) to calculate the confidence interval (CI) for the estimated LD50. This procedure is a statistical calculation that does not require the use of additional animals. The CI helps to place the estimated LD50 in a statistical context for hazard and risk assessment purposes.
- The U.S. EPA developed a software program for use in establishing test doses, determining when to stop the test, estimating the LD50, and providing a CI for the LD50. The publicly available software was developed to mitigate complexity for the user and to facilitate correct performance of the UDP.

The UDP Technical Task Force provided the following clarifications regarding animal welfare:

- The UDP guideline significantly reduces the number of animals used in comparison to OECD TG 401 by the incorporation of the following: 1) a stopping rule which limits the maximum number of animals in a test; and 2) a sequential dosing method which introduces further efficiencies in animal use.
- The UDP guideline provision that the initial starting dose should be below the LD50 will result in fewer animals receiving lethal doses, thereby providing further potential reduction in pain and distress.
- Adherence to the OECD Guidance Document on Humane Endpoints (2000a) should provide additional reduction or minimization of pain and distress in animals used in this procedure.

The revised version of the UDP and the UDP software program were then provided to the Panel and made available for public comment in July 2001(*Federal Register*, NIEHS, 2001a).

August 21, 2001 Peer Review Panel Meeting

The UDP Panel met, via public teleconference, on August 21, 2001 (*Federal Register*, NIEHS, 2001b). The agenda topic of the teleconference meeting was the scientific peer review evaluation of the following:

1. The revised draft UDP, modified in response to recommendations from the July 2000 Panel meeting;
2. A proposed procedure for calculating the confidence interval (CI) for the estimated LD50; and
3. A software program to aid in establishing test doses, determining when to stop the test, estimating the LD50, and providing a CI for the LD50.

The Panel was to evaluate the following:

1. The extent to which the revised draft UDP test guideline (July 12, 2001) incorporates modifications in accordance with the Panel's recommendations at the July 25, 2000 Peer Review Panel meeting;
2. The appropriateness and adequacy of the proposed procedure for calculating a CI for the LD50; and
3. The adequacy and consistency of the software program for use in the revised draft UDP test guideline.

Conclusions and recommendations from the Panel were as follows:

Revisions to the UDP Test Guideline

The Panel concluded many of the recommended and requested changes had been appropriately considered and all members concurred with the current modifications. However, several previous recommendations appeared to have not been

adequately addressed in the revised UDP Test Guideline, and the Panel recommended adding the following:

1. Either sex of animal can be used, or if information is available indicating that one sex is more sensitive, the more sensitive sex should be used.
2. A practicability evaluation of the usability of the *in vivo* test should be conducted to supplement the computational analyses.
3. A separate section on how the revised UDP Primary Test addresses reduction, refinement, and replacement of animals when compared to the previous tests should be included to the UDP guideline.
4. Constant concentration in dosing should be used unless there is a clear scientific or regulatory justification for using constant volume. In the event that constant volume is used, information on the actual concentrations utilized should be provided.
5. Additional guidance pertaining to the use of pre-start data (data available before the acute toxicity test is conducted) which may be helpful in determining the starting dose level should be provided.

Confidence Interval Procedure

Calculation of confidence intervals (CI) provides a basis for evaluating how to incorporate test results into regulatory applications. Therefore, a CI calculation was included in previous versions of the UDP guideline (OECD 1998 and ASTM 1998). Following deletion of the proposed supplemental procedure from the previous draft Revised UDP as per recommendation by the July 2000 Panel review, another method was needed to assist the investigator using the UDP to calculate a CI for the LD50. Based on this need, the U.S. EPA developed a proposed procedure for obtaining the CI; this procedure is a statistical calculation that does not require the use of test animals beyond what is needed to estimate the LD50. Further, the procedure helps to place the estimated LD50 in a statistical context for hazard and risk assessment purposes.

The Panel endorsed the proposed procedure for calculating the confidence interval for the estimated LD50. However, the Panel recommended the inclusion of language in the UDP guideline and software to fully describe the limitations and uncertainties of the proposed method, and to provide appropriate cautions for interpretation of test results. The Panel noted that statistical techniques are evolving and recommended the future development of alternative approaches, such as nonparametric methods, be encouraged.

Software Program

To support the modifications in the revised draft test guideline, a software program was designed and made publicly available to aid in the guideline procedures, to facilitate performance of the UDP, and to mitigate its complexity for the user. The U.S. EPA developed the Acute Oral Toxicity (U.S. EPA Revised Test Guideline 425) Statistical Program" (AOT425StatPgm) to perform the statistical calculations associated with the guideline. The AOT425StatPgm program performs the calculations required to complete the test procedure by calculating 1) the doses for the test animals, 2) when to stop dosing animals, and 3) the specified LD50 and a confidence interval for the LD50. Additionally, U.S. EPA conducted quality assurance testing and simulation testing to assess the performance of the software program and to determine the statistical performance of the OECD TG 425 procedure under various conditions.

The Panel concluded the software program was appropriate and suitable for establishing test doses, determining when to stop the test, estimating the LD50, and providing a CI for the LD50.